

---

# Badminton

## Navigate through badminton in France



Louis-Maxence GARRET & Karine PERRARD - 18 décembre 2018

---

**" Badminton is a sport which, when compared with football and tennis, still needs promotion. "**

**Lin Dan**



---

# Introduction

The goal of this process book is to show the reader the evolution of our project Badminton and explain the different phases we have been through, from the design creation to the final implementation.

This project was implemented in the scope of the EPFL course Data Visualization (COM-480). It uses some of the principles seen in that course, and follows the heuristics presented during the lectures. We tried to make our visualization clear and intuitive so that it can be used by any one, whether or not the user has bases in data visualization websites

## Overview

A very common habit for badminton players is to frequently check at their own statistics as well as the ones of others. This is due to the fact that being an individual sport, badminton is also very competitive, since competition is a way to assess one's level and progression. Competition is also a way for players to meet and share advices. Players are ranked using points and levels that relate their badminton skills as well as their competition involvement. In France there are more than 150'000 badminton players and 2'000 clubs, so we wanted to give this big community a way to notice each others.

## Motivation

For now, all of the french websites that show badminton statistics requires the users to know the name of the player or the name of the tournament. What is different in our project is that we want to let users explore the data, instead of searching in it. The idea is that we propose the data to the user in a graphical representation, and we develop it when they click on some points. To do so we used layers: we layered the players under the clubs, and the clubs under the geographical regions. That way one can explore without being overflowed with the information

We wanted to provide a way to players to check their statistics in an easy manner. The players can also search in the data to have a quick access to what they are interested in. With our visualization they should also be able to know where they can meet other players, for examples in which clubs and tournaments.

## Target Audience

We are targeting the badminton players in France, because the data that we have is from a french website, which regroups the informations about french players. We tried to keep our visualization as simple as possible, so that every badminton player can use it and understand it, no matter their age or their computer skills. The only requirement for having a good comprehension of this visualization is to have a basis knowledge in badminton, for example what do the levels mean.

---

## Related work and inspirations

One member in the group is a french badminton player, and knows well the website we used for data scraping. Thus the data exploration is a big problem from those websites (as said above), and we wanted to fix that.

We were also very inspired by the tweet map shown during the course, as its design was simple and powerful, and also easy to navigate in. (TweetMap from <https://www.omnisci.com/demos/tweetmap/>). We also got inspiration from all the websites that relates to badminton players in France, such that [www.badiste.fr](http://www.badiste.fr), [www.badmania.fr](http://www.badmania.fr), <https://poona.ffbad.org> ...

---

## DataSet

As stated before, there are many websites from where it is possible to fetch the data about players, clubs and tournaments. We were familiar with [www.badiste.fr](http://www.badiste.fr), and on this site some data can be directly loaded as .csv files, and the other data are stored in tables, which are practical to fetch from source code. Then we went through different phases to get the data we are actually using.

### Website analyzing

We took some time to go through the website and its data, as well as how it is built. To do so we make a big use of the inspect element tool provided in browsers and of the PostMan extension on Chrome, since we needed to intercept the POST requests made by the searches.

For tournaments scrapping it was quite straightforward, as all the tournaments are shown in the page <http://badiste.fr/liste-tournoi-badminton>. We needed additional informations, so we also analyzed the tournament page, it was built the same way for all tournaments (see page <http://badiste.fr/tournoi-badminton/tournoi-de-noel-du-bcbb-10021.html> for an example). We had to make sure that we fetch all tournaments, so the final link was <http://badiste.fr/liste-tournoi-badminton?c2=0&u2=1&page2=1#table2>, and we had to fetch that for the 8 pages (page2=1, 2...8) to get all the tournaments. We couldn't just download the provided tournament CSV because there were some tournaments missing and we needed some informations that were not present in the file.

For the clubs we first simply downloaded the file given in <http://badiste.fr/liste-club-badminton?sigle=&nom=&ville=&Submit=Rechercher> (under the Télécharger button at top right), but it appeared that some data were missing, so we needed to go through all club pages (ex: <http://badiste.fr/club-badminton/asptt-strasbourg-244.html>) to get that additional info.

The trickiest data to get was the players, since there is no page regrouping all the players. To get players we had to fetch players by club from the page <http://badiste.fr/rechercher-joueur-badminton>. When entering a club name the players are displayed at the bottom of the page. We yet had some issues, because we searched with the club nickname (ex: ASPTT for ASPTT Strasbourg), and several clubs have the same nickname. For those clubs we searched by nickname and by town, since no two club have the same nickname in the same city. For players we have all the informations we need for our current design in the search page, so we didn't need to go through all player pages.

Once we had a clear idea of how the data are shown it was possible to extract it using python code.

### Data extraction

All the processing was done with python code. We used the library [urllib](http://urllib) to scrap the page source codes. That returns the code as a string. Since it is HTML code, we can extract the different elements thanks to the library [BeautifulSoup](http://BeautifulSoup). To ensure that we did

---

not DDOS the website, we have a `time.sleep(0.5)` wait period between two page fetches. We store all the data extracted in a `pandas.DataFrame()` object.

For the clubs that we have as CSV, we were missing some infos. We got each club LatLong coordinates using python [geopy](#) library. We also get an HTML code snippet for each club, containing some infos such as website, mail and logo.

We had some issues fetching the whole data at once as the iterations became slower, so we fetch the data in batches and then we merged those batches, to save time and RAM memory.

## CSV creation and cleaning

With Dataframe it is very easy to create CSV files. Yet once we loaded the data in our project we found out that we still had some issued. For example the geopy library returned Vienna in Austria instead of Vienna in France. For those around 70 cases we manually changed to the right coordinates. An other example is that in Paris there were 20 clubs at the same position (Paris always have the same LatLong coordinate), so we got the correct coordinates from club websites.

---

# Exploratory

In this part we will explain what designs we considered. We will not talk about exploratory data analysis because we did not do it, since we scrapped the data by ourselves. That part was done when we were exploring the website from where we got the data. Since the beginning it was planned to do a map to show the data, but we will show all the steps through which we've been to get to the actual design.

## Designs considered

We have had a very top-down approach: for each part of our visualization, we brainstormed and implemented the most important ones first, and once we thought that the design was well-fitted to the data we switched to another implementation detail.

The first part we changed was how to display the datapoints. We've been through several designs that we will show and explain below:

### 1) Graphical elements

#### a) zoom

1: continuous slider

At first we considered having a slider that controls the zoom level, which we have in many maps.



2: buttons + France level

Then we figured out that it would be less space-consuming to have buttons. That allowed us to have an option to go back to the country level. Thus our final design with the points/density changes occurring with zoom is better-fitted for a discrete zoom rather than a continuous one.



#### b) legend

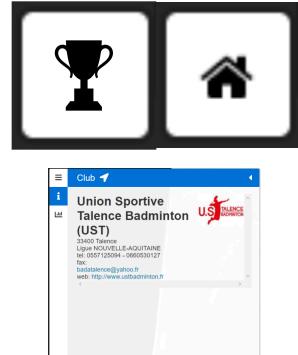
At the beginning we didn't have the need of a legend. But then when the density color feature appeared we displayed one. We have never changed it.

---

### c) sidebar

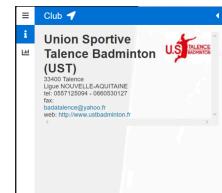
#### 1: buttons

We used to have one button to switch between clubs and tournaments, but that didn't offer the possibility to display much informations.



#### 2: sidebar

To be able to show more informations and to propose some new functionalities we decided to create a sidebar that one can hide and show.



### d) popups

When we click on or hover a datapoint we have the possibility to display a popup window. The contents of that window has had a lot of changes:

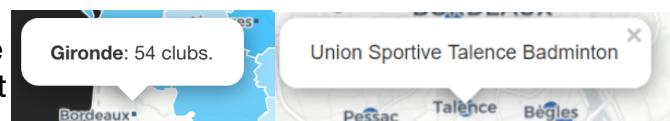
#### 1: club infos

In our initial design we wanted to show club informations showing in a popup when a datapoint was clicked. That design didn't allow to display as much informations as we would have wanted, so we decided to put that in a sidebar.



#### 2: department density and club name

To not loose the popup functionality we decided to still display something. When the zoom is at country level it displays the actual density on department hover. But at department zoom level one needs to click on a point to show its name.



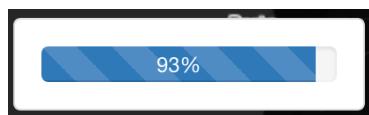
### e) description box

We used to have a description box, but with the design evolution it finally contained only redundant information, so we removed it



### f) loading bar

The data may be quite slow to load, so to let the user know that the visualization is not crashed we show a loading bar when there are no more informations to display.

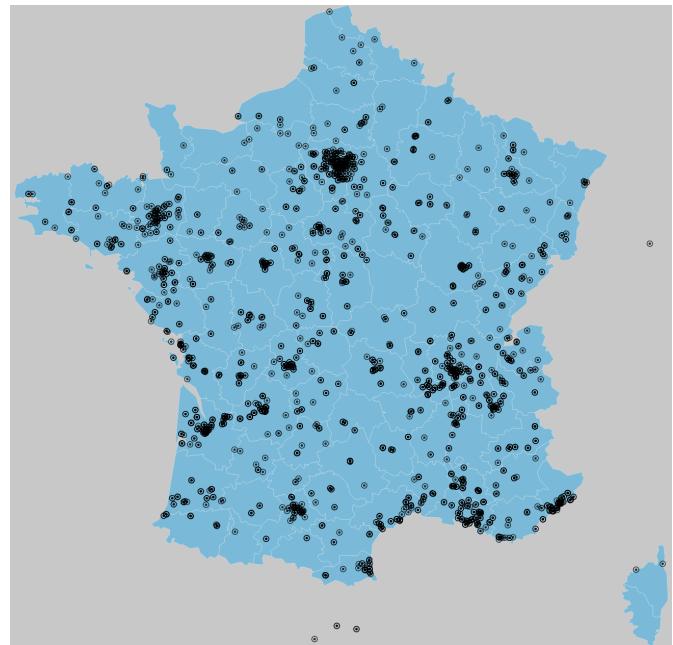


---

## 2) Datapoints

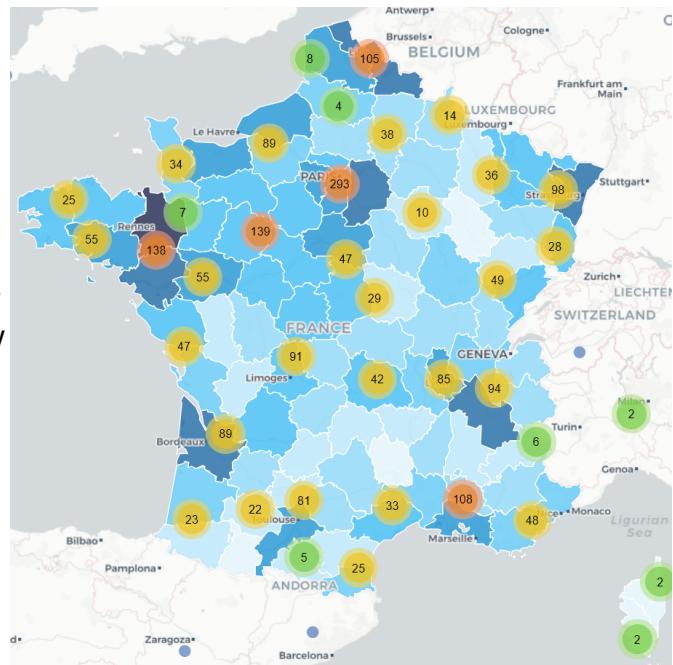
### 1: Showing all the points

We got that idea from the tweet map we were inspiring from, but once it was implemented it appeared to us that is created a visual clutter, and we needed to find a way to fix that.



### 2: Showing points clusters

To fix the visual cluttering, we decided to group the points from a department into clusters, and update the clusters relatively to the zoom level. Thus there were too few informations with that design, we were loosing a lot of space that we could have used to show information

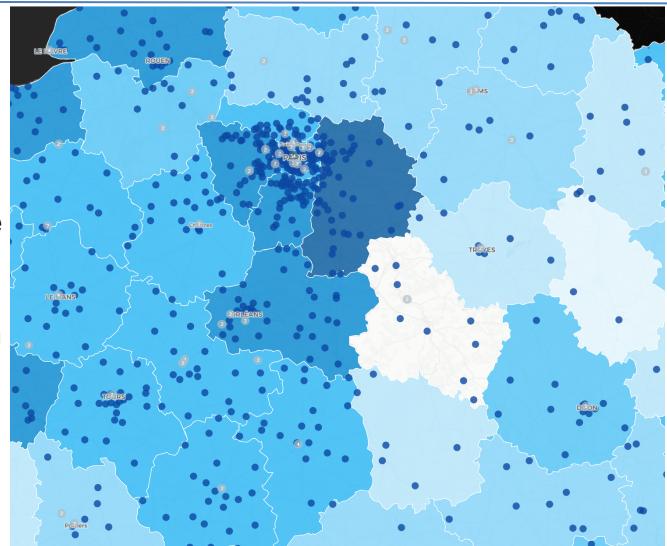


### 3: showing clusters and density

The good point of showing all the points was that we could have an overview of the point density over the map. To include this information back in our design we changed the department color based on how many data points there are in this area. We still kept the clusters. But yet there was information redundancy.

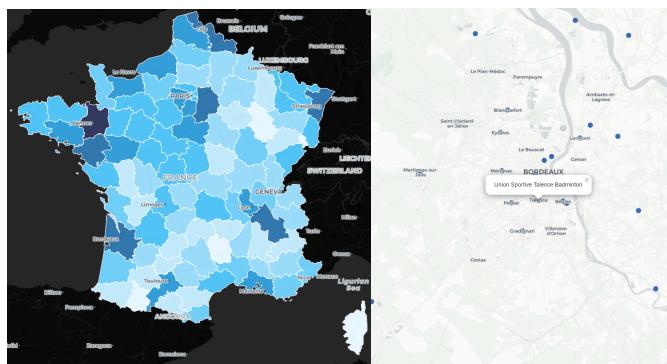
4: showing density, then density + points at a given zoom level

To solve that redundancy problem, we decided to not show the clusters when the zoom is at its minimum. The points are shown when we arrive at a given zoom level where it does not produce visual cluttering, and the clusters are now shown only when several points are at the exact same position. But we still show density at any zoom level.



5: showing density, then points at a given level

To avoid any information redundancy, we don't show the color (density) when the user clicks on a department and that the zoom is focusing only on it. That allows the user to focus and that department and its datapoints.

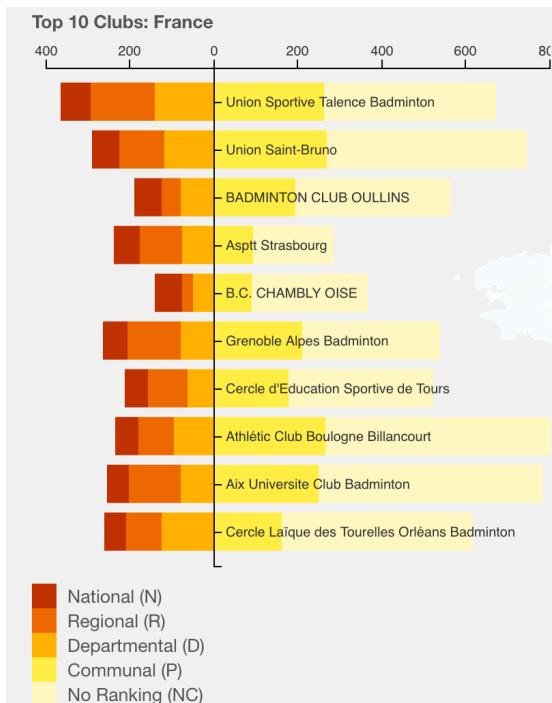


### 3) Sidebar Informations

a) ranking

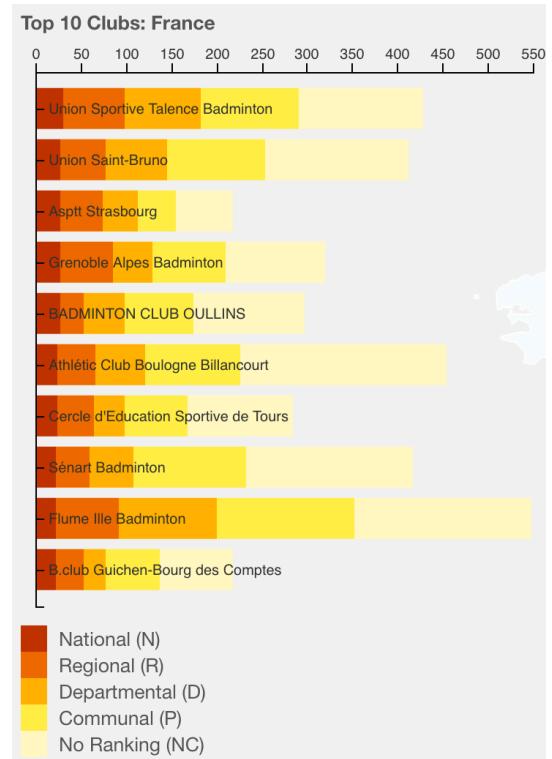
### 1: diverging stack bar

We wanted to do a diverging stack bar to represent the amount of players above and under a given level for each club. But the problem was that we have amounts on both sides, so we cannot have negative and positive numbers.



### 2: stack bar

The easiest way to show the amount of players per level in a club is a bar chart. Since we don't want our design to create visual cluttering we decided to do a stack bar, that represents all the information needed on one bar.

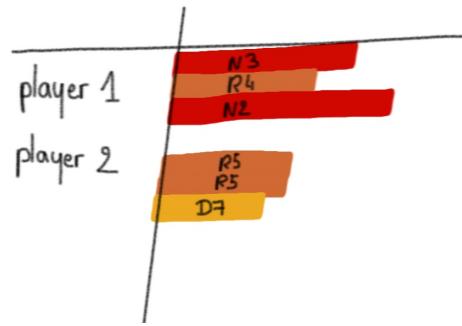


### b) clubs

We first thought about displaying just the club informations, but for more practicality we added all the clubs players in the page, and a link to go to the player page. We have considered several representations to show the players:

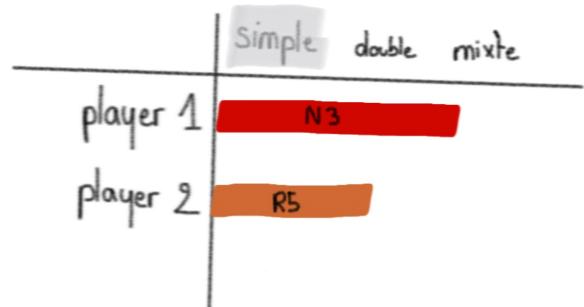
### 1: bar chart

Since players have levels in 3 categories we wanted to show them all at once to have an



### 2: interactive bar chart

Then we considered to show the bar chart for one level at once, and provide a way to



### 3: table

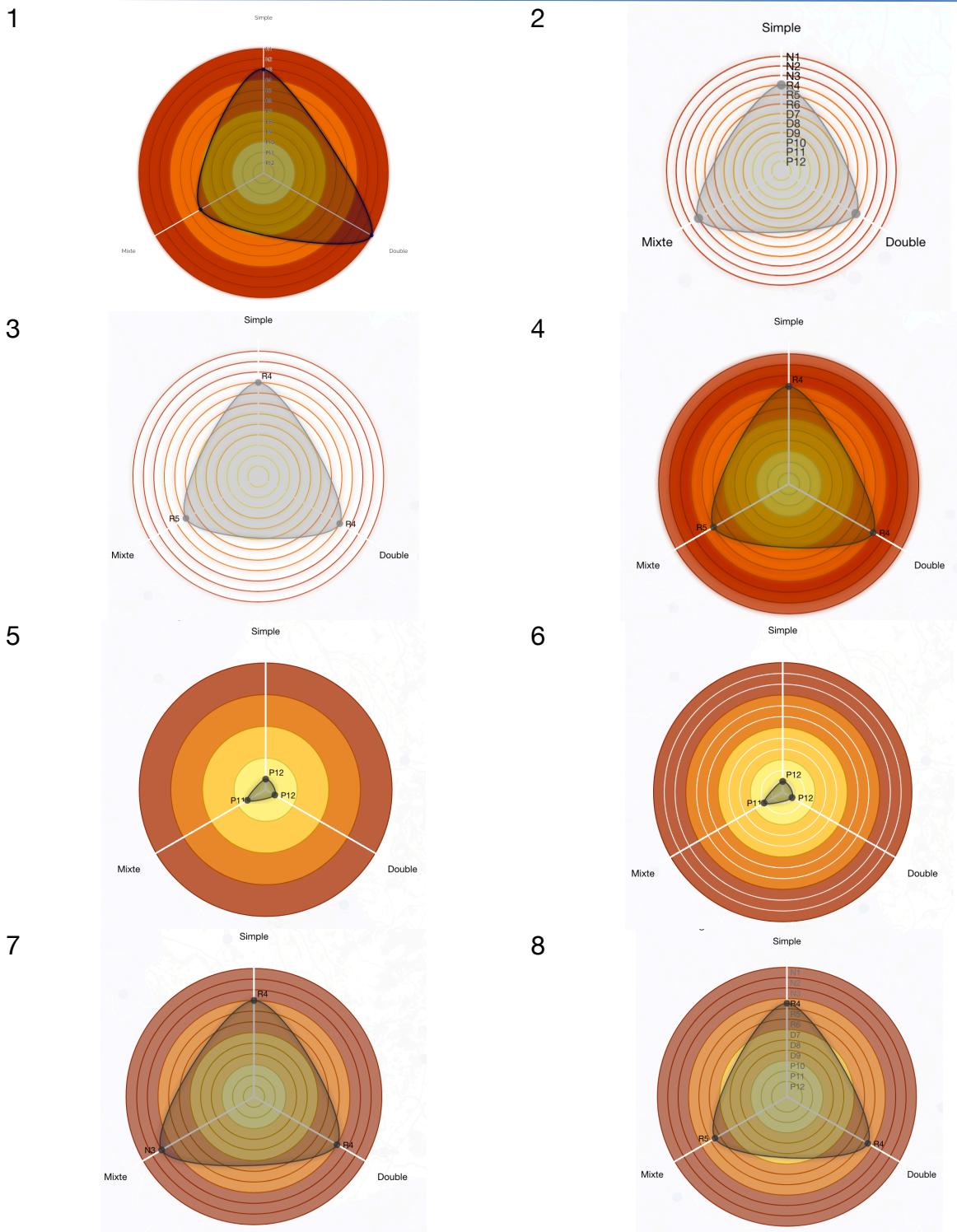
Since the player levels are values, we can just display it as a table. To highlight the data, we

### Les joueurs en 2018-2019

Nom	Prénom	Genre	S	D	M	Moy
Lucas	CLAERBOUT	M	N1	N2	N2	6495.13
Jordan	CORVEE	M	N2	N1	N1	5820.36
Anastasia	CHERVYAKOVA	F	N2	N1	N1	5500

### c) players

We create a page for players that displays all the informations we have. We decided to present the player level as a star chart to highlight it, since the goal of our visualization is to show the levels of the players. We have considered several design for that star chart, but after brainstorming it appeared that the last one was the best one, since the level colors are displayed but not too vibrant, and the legend is displayed as well as the levels. We showed the values of each point because it was tricky to follow the lines up to where the legend was placed in the first design.



---

# Implementation

## Functionalities

When the page is loaded the visualization displayed is the map, at the scale of France. There is also the side bar that opens, to give the user some informations about the project and how to use it. Then from that point you have several interaction options.

### Mouse hover

At the France zoom level there is the possibility to get some informations on mouse hover. When the mouse is on a department there is a popup window that appears, showing what is the exact number of club in that given department. The color and the legend provide a preview of that data, but it doesn't show the exact amount of clubs.

### Zoom

Like most maps we provide a zoom feature that allows the user to zoom in and out the map. Zooming out allows to see some overseas departments. Zooming in allows to have more detailed informations about a geographical region. Our visualization evolves while the user zooms in. At first the departments are colored according to their club density, but when the user zooms on a department (that can be done by clicking on the department) the color disappears and the data points appear. Showing the points only from a given zoom level allows to still provide information without cluttering. The user can then click on the points to get informations about them.

### Club ranking

When the user clicks on the ranking tab of the side bar there is a ranking of the clubs that is displayed. The clubs considered are the ones present in the map: if the user is at the France level then the best 10 french clubs are shown, and when a department is clicked that ranking is updated with the best 10 clubs of that department.

### Datapoint interaction

When the user clicks on a datapoint the sidebar gets opened on the information tab. In that tab all the informations of the clubs are shown, with many clickable links, for example club website. In the table underneath all the players from that club are shown. They are initially sorted by mean point score, but it is possible to change that sorting by clicking on the table head (like most tables). The users are also clickable, and when doing so the info pane changes into the informations of the player instead of the informations of the club.

### Player display

The player page doesn't provide much interactions, is mainly shows informations. The user can go back to the club page by clicking on the "<Go back to club" button at the top.

## Search

Even though our vis is mostly about exploring, we still wanted to keep the search feature to not loose the usability of the websites we got inspired from. In order to search the user needs to enter at least two characters, and the search is interactive, i.e the search results are updated upon each new letter. The players are flagged with a user logo, the clubs with a location logo.

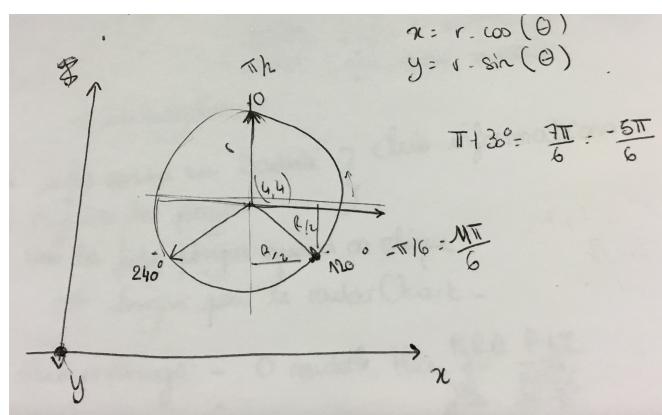
## Technical details

For most of the implementations we used libraries and online tutorials. To adapt them to our visualization we made sure to fully understand the code and then to change the right values, remove useless parts and add what we need.

The first example we used is for the map. We actually took several examples given in <https://leafletjs.com/examples.html>. Of course we started with the base tutorial <https://leafletjs.com/examples/quick-start/>, then used <https://github.com/Leaflet/Leaflet.markercluster> in order to group our points into clusters. Based on this example <https://leafletjs.com/examples/choropleth/> we were able to display the department map as wanted. After that, many of our modifications were internal code changes.

Above that we needed to find the right way to display the informations through a side bar. Once we had the side bar design we used the sidebar-v2 library <https://github.com/nickpeihl/leaflet-sidebar-v2> for which we had to add our custom methods in order to interact with and update panels in the sidebar. Thus this library's code is included and modified in our repository.

We also created several charts, such as a bar chart and a star chart. For the star chart we used the examples given in <http://blocks.org/mthh/7e17b680b35b83b49f1c22a3613bd89f>. We removed the interactions parts from the code, and we changed the circles colors and the legend. There was some trigonometry implied for the display of the points values, as shown in the figure below.



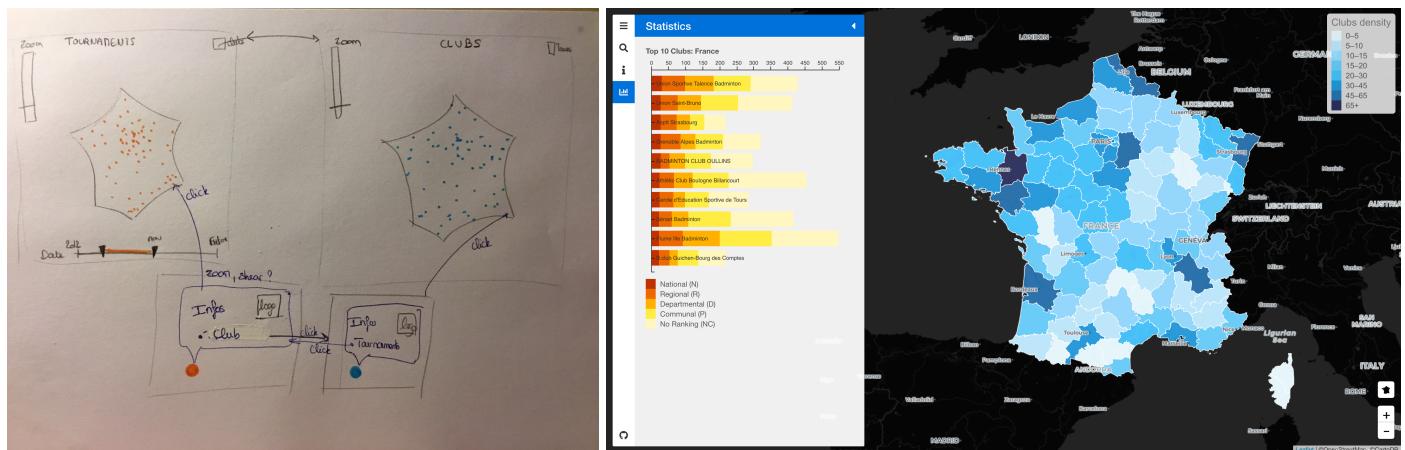
For the stack chart we used the code given in <https://gist.github.com/hydrosquall/88b1511b74faee969196c93e25cc3890>. At first, we wanted to display the repartition of the players' levels in the club using a diverging stack bar. As it was not clear and did not make sense to define a 0 "between" players. Thus, we changed from this idea to a simple

stack chart, were the comparing levels (e.g. mostly National) are first in the stack bar. This allows us to sort clubs depending on this first level and visually see the difference quickly.

## Divergences from initial model

The biggest difference is that in our final model we don't present the tournaments. There are several reasons for that. First we didn't figured out how to include them in an intuitive way in the visualization. We thought about adding a button but there there would be no link at all with the current club and players data. The second reason is that the processing time for fetching and cleaning the data was very long, and we didn't have the time to get the data linking the players to the tournaments, as well as the results of the players. We finally decided that it was better to not include the tournaments in any way rather than having a button or a tab that is not linked to anything.

Below you can see the sketch for our initial model and the final model. It clearly shows that we stuck to our initial idea of having a map, but that the surrounding features went through a lot of thinking and modifications to get to what they are now.



---

## Evaluation

Now we will show a resume about what we learned from that project, and what we think are the strengths and the weaknesses of our model. The goal of that project was to provide a tool for players to explore badminton in France, and we think that this model actually provides a good base for that purpose. Yet it can still be improved.

### Data insights

During this whole project we have been working on the same data, so we won some knowledge on it. The first thing that we really didn't imagine is that the clubs repartition in France is very disparate. In some regions there are a lot of clubs, sometimes several are in the same little town, whereas in the region next to if there are only a few, separated by dozens of kilometer. One could think that in a country as developed as France there would be no such inequalities, or at least not that flagrant. Indeed the geographical repartition of clubs reflects more the urbanization phenomena in France rather than a lack of interest in badminton, more generally in sport, in regions of the country.

We also noticed that the clubs where the best players are are mostly in big cities. Thus those clubs are containing a lot of players, whatever their level. We could make several hypotheses about that. The first one would be that the clubs with good players actually attracts the other good players, since they see there an opportunity to improve their badminton skills, and also a way to confront themselves with other good players. An other hypothesis is that the repartition of player levels in clubs is constant, but since that there are more players in some clubs then inevitably there are also more good players. Unfortunately we don't have a way to verify our first hypothesis. The second one could be checked if in the ranking we show percentages instead of numbers. We didn't want to do that because we want to show the club magnitude.

### Critiques

We want to let the users discover what is going on with the badminton, and we achieved that. Yes the users might be a little too much left to themselves. It is true that when the page is loaded, if the user don't have a specific goal (such as check clubs in Isere for example) then the model may seems quite superfluous. Maybe some gamification

Thus the visualization may be quite slow, due to all the data loading and processing. That might be a limitation for user experience, since nowadays models are expected to be fast and smooth, but we lack smoothness in our model.

Another point is that we spent quite some time analyzing the website where we got the data from, and also some time fetching those data. This is time that we lost in working on the model, maybe it could have had some more features with a little more time.

---

## Further improvements

We believe that our model could be make much more powerful with some more work and data. The first thing that is in our opinion the most important to improve is the smoothness of the model. It can be discouraging for users to have a slow and batched model. It can make their experience unpleasant, and they wouldn't be willing to use our visualization anymore. The goal of our project is to provide experience and help, not pain and hardness.

Another point that we would have loved to have is tournaments. The issue here is that we were missing some data that would have required too much time to process, and se we did not know how to include the few data we have in our design. We wanted to have an interactions between the players, the clubs and the tournaments but unfortunately that could not be done in the scope of this project. It would still be very a very interesting improvement to add a nice interaction with players and tournaments and clubs.

The data handling could also be improved. The way we grab the data is not the best, maybe contacting the french badminton federation to get full data is a great idea. Then we could work with a very complete and rich dataset that wouldn't limit the model in any way. The other point es that we do not have a database, we are roughly using the csv files because our data is small enough to do so. Using a database may be good to provide a model that can be extended to more data (maybe more countries).

The last point is that it would be very welcomed to add more interactions and animations, as long as that adds meaning and usability to the model.

---

## Peer review

### Review of Louis-Maxence

Preparation: We mostly communicated by messages but still had informal meetings on Tuesday. At the end of each meetings that we had during the labs on Tuesday we held a TODO list of what we wanted to see implemented, experiment with or change. This way, the preparation was implicit for the next meeting, and we would update our task list during the week so that we knew where we were when meeting again.

Contribution: The workload was quite consequent as we were only two in the team. Naturally, we split the tasks based on the task list equally. This might even be easier when there are two people than three to correctly split the workload! Throughout the semester, it happened that Karine mostly worked on the scraping and the data processing, when I worked mostly on the visualization, even though in the end I had a few times to work on the data and she had to work on the visualization as well. In terms of work hours the contribution is equal during the semester, even though the balance during a single could hugely vary, as we both had other demanding projects with different deadlines. In the end, we managed this aspect of project management as quite well!

Respect for others' ideas: We had many times difficulties to understand each other points of view, hence man of the sketches drawn were mostly to get the other part to understand ones' idea. Even though we worked on different parts of the project frequently, we still asked the other's opinion about the model we wanted to build. Of course there were some disagreements, hence often in this case we would ask the opinion of a third person.

Flexibility: Sometimes it was hard to consider the other's idea if we were already very advanced in the implementation (because of course no one likes to discard ones work) but by discussing first of what we wanted to do and make sure we were on the same page, this situation did not occur much. Otherwise, we had no issue with showing disagreement with the other's idea and have a constructive talk about it. Many parts of the projects come from a disagreement that lead us to finding an idea that seemed good for both of us!

We had already worked on other projects together and had no issue work again together, as we knew that our personal way to work in projects were compatible. Teaming up with Karine for this projects was a good experience again.

### Review of Karine

For me the preparation was all good, even if we didn't have proper meetings. Since we were only two in the team then a lot of the discussions occurred by messages. We met every Tuesday during the lab session, and it seems that during this meeting both of us were knowing what we talked about. It sometimes happened that one of us was lost, but the reason was that the other was not clear, and in that kind of situation the other reformulates so that we both understand.

In my opinion we both contributed to build the model in a similar way. I've been working a little bit more on the website analyzing and on the data processing, but Louis-

---

Maxence compensated by working more on the project implementation, so that the overall work seems quite equilibrated to me. For the visualization design we both proposed our own ideas, and we tried to make it work so that we could include as much of our ideas as possible in our model.

The flexibility was sometimes more tricky, but that was the case for both of us. I know that I might also not be flexible some times. Still it was never about big disagreements, we always agreed to argue. When that kind of situations occurred, we both sketched our idea, the pros and the cons and showed it to the other. Sometimes we disagreed. The problem was not that we were stuck to our model or that we refused other's, but rather that even after the explanations we still didn't see what benefits one model brings to the model more than the other. In case of disagreement about what we should do we asked to the teacher or to other student what they think. It happened that their answers were often a mix of our both proposals.

In general when one of us exposes his ideas the other listens and comments it, it can be either critiques, agreements or other ideas that can be added to that proposal.

I have already had the opportunity to work on a project with Louis-Maxence, and we know that our duet is working well and that we both are open to discussion. That project was not an exception and I had a lot of pleasure to work with him on that project.